Information causality in multipartite scenarios

Lucas Pollyceno ,[1,*] Rafael Chaves,[2,3] and Rafael Rabelo [1]
[1]*Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas, CEP 13083-859 Campinas, Brazil*
[2]*International Institute of Physics, Federal University of Rio Grande do Norte, 59070-405 Natal, Brazil*
[3]*School of Science and Technology, Federal University of Rio Grande do Norte, 59078-970 Natal, Brazil*

(Received 18 January 2023; accepted 20 March 2023; published 7 April 2023)

The Bell nonlocality is one of the most intriguing and counterintuitive phenomena displayed by quantum systems. Interestingly, such stronger-than-classical quantum correlations are somehow constrained, and one important question to the foundations of quantum theory is whether there is a physical, operational principle responsible for those constraints. One candidate is the information causality principle, which, in some particular cases, is proven to hold for quantum systems and to be violated by stronger-than-quantum correlations. In multipartite scenarios, though, it is known that the original formulation of the information causality principle fails to detect even extremal stronger-than-quantum correlations, thus suggesting that a genuinely multipartite formulation of the principle is necessary. In this work, we advance towards this goal, reporting a different formulation of the information causality principle in multipartite scenarios. By proposing a change of perspective, we obtain multipartite informational inequalities that work as necessary criteria for the principle to hold. We prove that such inequalities hold for all quantum resources and forbid some stronger-than-quantum ones. Finally, we show that our approach can be strengthened if multiple copies of the resource are available, or, counterintuitively, if noisy communication channels are employed.

DOI: 10.1103/PhysRevA.107.042203

I. INTRODUCTION

It is undeniable that quantum theory is one of the most successful scientific theories ever developed. Its mathematical formalism and axioms, although counterintuitive, led to extremely precise and, oftentimes, intriguing predictions, which were confirmed experiment after experiment. Among the most counterintuitive phenomena that quantum systems can display, Bell nonlocality [1,2] is one of the most fascinating.

Bell nonlocality refers to stronger-than-classical correlations between outcomes of measurements performed on space-like separated systems. However, as noted by Tsirelson [3], and later by Popescu and Rohrlich [4], nonlocal correlations displayed by quantum systems are limited and are not as strong as they could be, in principle. From a formal viewpoint, such limitations follow from the mathematical axioms of quantum theory. From a physical viewpoint, though, it would be of interest to the foundations of quantum theory to identify operational axioms that could explain the limits of quantum nonlocality and, ultimately, lead to the derivation of the mathematical axioms of the theory from first principles.

In the last couple of decades, several principles were proposed with the goal to explain why quantum theory is not more nonlocal. Among them are the principle of nontrivial communication complexity [5], the principle of macroscopic locality [6], and the principle of local orthogonality [7]. Although all of the cited candidate principles are very good at identifying and forbidding unreasonable consequences of stronger-than-quantum correlations, they are provably not capable of excluding *all* stronger-than-quantum correlations [8].

One candidate principle that may be capable of singling out the nonlocal correlations allowed by quantum theory from more nonlocal ones is the information causality (IC) principle [9]. Roughly speaking, the principle states that, in a communication scenario, the receiver's available information concerning the sender's initial set of data cannot exceed the message information amount. The violation of the principle would allow, for instance, one to receive an amount of information corresponding to one page of a book, and choose, afterwards, which page the receiver would like to read.

It is well known that all quantum correlations, despite nonlocal, obey the information causality principle. It is also known that *some* stronger-than-quantum correlations violate the principle [9]. It is unclear, though, if *all* stronger-than-quantum correlations violate it. The main reason for this is the fact that the principle, although relatively intuitive, is very hard to formalize in the form of a mathematical criteria. A sufficient criterion for the violation of the principle was proposed in Ref. [9], but it was later proved that there were nonquantum correlations that do not violate it [10]. Since then, other techniques were developed to generate stronger criteria for the information causality principle [11,12], but, so far, none of them has shown to be strong enough to characterize the exact set of quantum correlations, even in simple scenarios.

Another requirement that was discovered recently is that any operational principle has to be genuinely multipartite to correctly retrieve all quantum nonlocal correlations [13]. This observation complements the finding that the bipartite formulation of the original IC proposal cannot be applied to

*lpolly@ifi.unicamp.br

exclude even some extremal tripartite stronger-than-quantum correlations [14]. Hence, it is clear that a genuinely multipartite formulation of information causality is necessary for it to be a valid and tight operational principle for quantum theory.

In this work, we propose a multipartite perspective for the information causality principle. First, we establish a multipartite communication task that, in a sense, generalizes the random access code (RAC) task associated with the original IC formulation of Ref. [9]. Thus, we present different multipartite informational-theoretic criteria, which ensure IC in the proposed scenario. We then prove that the inequalities hold for all quantum nonlocal resources and are violated by some stronger-than-quantum ones. Also, if many copies of the nonlocal resource are available, we show that, by applying the concatenation approach presented in Ref. [9], and in analogy to the result obtained in such a reference, one obtains stronger criteria for IC. In addition, we show that our findings are in agreement with the recent results reported in Ref. [15], where it was shown that the employment of noisy communication channels leads to the same constraints as the concatenation procedure mentioned.

II. BIPARTITE INFORMATION CAUSALITY

The *information causality* scenario [9], considers a strictly bipartite communication task: Alice encodes a bit-string $\mathbf{x}$ of length n in a message M of m bits (where $m < n$) and sends this message to Bob; he then decodes the message and produces a guess G_i about a randomly selected bit X_i (the ith bit of the string $\mathbf{x}$) of Alice. Within this context, the information causality principle states that *Bob's information gain about the initial n bits of Alice, considering all their possible local as well as preestablished shared resources, cannot be greater than the number of bits m, sent by Alice.* The principle, proven to hold quantum mechanically, is mathematically captured by an information inequality, written in terms of Shannon entropies as

$$\sum_{i=1}^{n} I(X_i : G_i) \leqslant H(M), \tag{1}$$

where $H(M) = -\sum_m p(m)\log_2 p(m)$ is the Shannon entropy of the random variable M described by the probability distribution $p(m) = p(M = m)$. In turn, $I(X_i : G_i) = H(X_i) + H(G_i) - H(X_i, G_i)$ stands for the mutual information between Alice's bit X_i and Bob's guess G_i. Put differently, as expressed in Eq. (1) above, the IC principle states that if Bob is trying to guess one of the random variables X_i in possession of Alice, the total accuracy of his guess, quantified by the sum of the mutual informations $I(X_i : G_i)$, will always be limited by the amount of information sent from Alice to Bob [as quantified by the Shannon entropy $H(M)$ of the message].

The information causality inequality (1) can be violated by postquantum correlations, that is, correlations incompatible with the quantum mechanical rules. A paradigmatic example is that of a Popescu-Rohrlich (PR) box [4], a nonsignalling (NS) correlation described by the probability distribution

$$p(a, b|x, y) = \begin{cases} 1/2 & \text{if } a \oplus b = xy, \\ 0 & \text{else.} \end{cases} \tag{2}$$

To verify that is indeed the case, it is sufficient to consider Alice has two input bits, described by the random variables X_1 and X_2. Thus, consider the following protocol in the IC scenario. Alice inputs $x_1 \oplus x_2$ on her share of the PR box, obtaining outcome a that is then encoded in the message $m = a \oplus x_1$ to Bob. Bob inputs $y = 0$ if his aim is to guess the bit x_1 and $y = 1$ if he wants to guess the x_2 of Alice, using as his guess $g_i = m \oplus b$. Using the definition of the PR box, if $y = 0$ we see that $g_1 = x_1 \oplus a \oplus b = x_1$. If $y = 1$, $g_2 = x_1 \oplus a \oplus b = x_1 \oplus x_1 \oplus x_2 = x_2$, implying that $I(X_1 : G_1) = I(X_2 : G_2) = H(M) = 1$, thus violating the IC inequality (1). Here and throughout the paper, we denote random variables by capital letters such as X_i and G_i and use the common shorthand notation $p(x) = p(X = x)$ to denote the probability of the random variable X taking value x.

Notice that the simple protocol above employs a single copy of the distribution $p(a, b|x, y)$, which operationally can be understood as a black-box taking local inputs and producing correlated outputs. By introducing a concatenation procedure (a version of which will be detailed below), one can also consider the case of multiple copies of identical binary-input and binary-output nonsignaling boxes. In particular, the concatenation procedure introduced in Ref. [9] shows that the IC inequality implies another constraint for nonsignaling correlations, given by

$$E_I^2 + E_{II}^2 \leqslant 1, \tag{3}$$

where $E_j = 2P_j - 1$ is defined in terms of the conditional probabilities $p(a, b|x, y)$ as

$$P_I = \tfrac{1}{2}[p(a \oplus b = 0|0, 0) + p(a \oplus b = 0|1, 0)], \tag{4a}$$

$$P_{II} = \tfrac{1}{2}[p(a \oplus b = 0|0, 1) + p(a \oplus b = 1|1, 1)]. \tag{4b}$$

In fact, this constraint is equivalent to the bipartite quadratic Bell inequality, the so-called Uffink's inequality [16] (interestingly, however, in the next sections we will show that, for multipartite scenarios, such equivalence no longer holds).

Of particular relevance is the fact that this mapping from the IC inequality (1) to Uffink's inequality (3) proves that any correlation beyond the Tsirelson's limit for the Clauser-Horne-Shimony-Holtz (CHSH) inequality [17] will violate the information causality principle and thus witness its incompatibility with quantum theory. More precisely, as proven by Tsirelson [3], the classically valid CHSH inequality

$$\text{CHSH} = \langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leqslant 2, \tag{5}$$

achieves a maximum value in quantum theory of $\text{CHSH}_Q = 2\sqrt{2}$. A PR box, in turn, leads to $\text{CHSH}_{\text{NS}} = 4$. A direct analysis shows that any distribution achieving $\text{CHSH} > \text{CHSH}_Q$ violates Uffink's inequality and thus has its postquantum nature witnessed by the information causality principle. One should remark, however, that it remains unclear whether the entire set of quantum correlations can be recovered from the IC principle [8,10,18]. Interestingly, there are known postquantum correlations known to not violate Uffink's inequality [11,19].

Motivated by that insufficiency of the standard formulation of the IC principle, a general informational-geometric approach was introduced in Ref. [11] and shown to lead to stronger IC information inequalities. For example, the inequality given by

$$\sum_{i=1}^{n} I(X_i : G_i, M) + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M) \leqslant H(M) + \sum_{i=2}^{n} H(X_i) - H(X_1, \dots, X_n). \quad (6)$$

The original IC inequality (1) is a particular case of Eq. (6). As a matter of fact, the stronger IC inequality (6) was proven, in some cases, to be even stronger than Uffink's inequality (3). More precisely, with a single copy, inequality (6) can detect the postquantumness of correlations that cannot be detected by Eq. (3) even in the asymptotic regime of infinite copies of the correlation under test.

Building upon the original IC criterion, a more recent approach [15], generalized it to consider noisy channels between Alice and Bob, namely,

$$\sum_{i=1}^{n} I(X_i : G_i) \leqslant C, \quad (7)$$

where $C \equiv I(M : M')$ is the noisy channel capacity defined in terms of Shannon's mutual information and M' is the message that reaches the receiver after passing through the channel. Contrary to the original formulation, this approach proposes to search for the strongest nonsignaling correlations allowed by IC for every possible noisy channel between the parts. A generalization that allows recovering standard results, for instance, the Tsirelson's bound implied by Uffink's inequality (3), most importantly, however, without the need of a concatenation procedure, that is, in the single-copy level. This stems from the fact that, as the channel's noise is increased, the bound in Eq. (7) becomes stronger, capable of detecting the postquantum nature of correlations that cannot be witnessed in the noise-free version.

The main goal of this paper, as will be detailed in the following, is to generalize all such results, stated so far only for the bipartite scenario, into the multipartite setting.

III. NEW CRITERIA FOR MULTIPARTITE INFORMATION CAUSALITY

Our first goal is to introduce a natural generalization of the bipartite IC inequality to the multipartite scenario. For that, we will closely follow the information-theoretic approach based on quantum causal structures envisioned in Ref. [11]. In this approach, information principles such as information causality are nothing else than entropic constraints arising from imposing a quantum description to a given causal structure. As such, each quantum causal structure will have associated with it a given set of entropic inequalities, each of which can be interpreted as an information-theoretical principle.

In this work, we consider a particular class of quantum causal structures that naturally generalize the known bipartite scenario. Consider N parts, among which $N-1$ are senders in possession of their respective bit-strings $\mathbf{x}^k = (X_1^k, X_2^k, \dots, X_n^k)$, where $k \in \{1, 2, \dots, N-1\}$. Each sender encodes a classical message M_k of size $m < n$ to the Nth part, the receiver who has to compute one out of n possible bits functions $f_j(X_j^1, X_j^2, \dots, X_j^{N-1})$, by producing the guess G_j, where $j \in \{1, \dots, n\}$. This scenario is illustrated for the tripartite case, as a directed acyclic graph (DAG), in Fig. 1. As proven in Appendix B, considering that, additionally to any local operations, every part may explore their preestablished correlations mediated by a joint quantum state ρ, the following multipartite version of information causality holds:

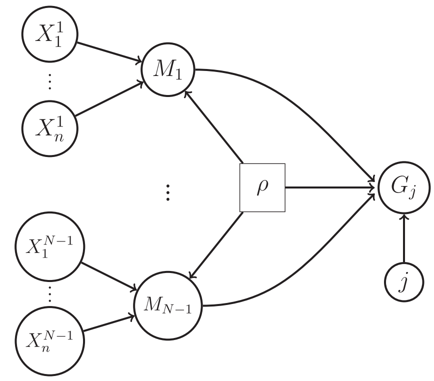

FIG. 1. Quantum causal structure, described as a DAG, associated with the multipartite information causality scenario.

$$\sum_{k}^{N-1} \sum_{i}^{n} I\left(X_i^k : X_i^1, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\right) \leqslant H(M_1, \dots, M_{N-1}) + \sum_{k}^{N-1} \sum_{i}^{n} I\left(X_{i+1}^k, \dots, X_n^k : X_i^k\right). \quad (8)$$

Or, in other words, the new multipartite inequality (8) holds for all correlations provided by quantum theory. To illustrate, we consider the tripartite scenario, depicted in Fig. 1, such that Alice and Bob have just two initial uncorrelated bits and that the communication task of Charlie is to compute two specific functions $f_1 = x_1^1 \oplus x_1^2$ and $f_2 = x_2^1 \oplus x_2^2$. The communication task is trivialized if the parties share the following tripartite nonsignaling (postquantum) correlation [20]:,

$$p(a, b, c|x, y, z) = \begin{cases} 1/4 & \text{if } a \oplus b \oplus c = xz \oplus yz, \\ 0 & \text{else}, \end{cases} \quad (9)$$

where $a, b, c, x, y, z \in \{0, 1\}$. To achieve it, the parties perform the protocol detailed in Fig. 2. In each run, Charlie can always perfectly compute each of the functions, since $g_1 = x_1^1 \oplus x_1^2$ and $g_2 = x_2^1 \oplus x_2^2$. In other words, similarly to the usual information causality scenario, Charlie has potential access to the four bits of Alice and Bob but receives just two bits communicated by them. Particularizing inequality (8) to

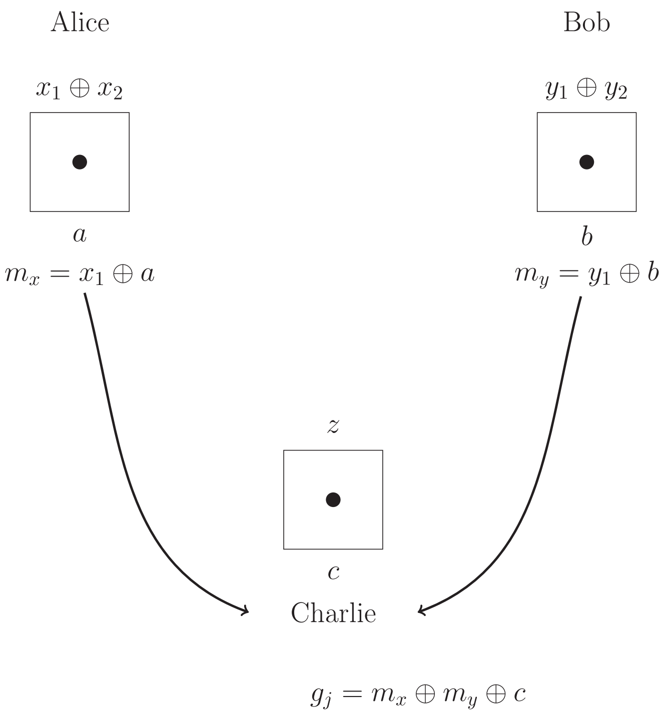

FIG. 2. The communication protocol is performed by Alice, Bob, and Charlie that share a nonsignaling resource. Alice (Bob) receive initially two bits $\{x_1, x_2\}$ ($\{y_1, y_2\}$) and perform her local measurements as $x = x_1 \oplus x_2$ ($y = y_1 \oplus y_2$). After obtaining her outputs a (b), Alice encodes the message with $m_x = a \oplus x_1$ ($m_y = b \oplus y_1$). Charlie inputs on his side $z = 0$ if he wants to compute f_1, and $z = 1$ if he wants to compute f_2. After receiving the messages, Charlie computes his guess by following $g_j = m_x \oplus m_y \oplus c$.

this case, we obtain

$$\begin{aligned}\mathcal{I} = {} & I\left(X_1^1 : X_1^2, G_1\right) + I\left(X_2^1 : X_2^2, G_2\right) \\ & + I\left(X_1^2 : X_1^1, G_1\right) + I\left(X_2^2 : X_2^1, G_2\right) \leqslant H(M_1, M_2),\end{aligned} \tag{10}$$

an inequality that is maximally violated by the NS correlation (9) with the protocol described above since $\mathcal{I} = 4$ while the quantum valid upper bound is $H(M_x, M_y) = 2$. In fact, the multipartite version (8) can be violated by the multipartite extension of the postquantum correlation (9), given by

$$\begin{aligned}& p(a_1, a_2, \dots, a_N | x_1, x_2, \dots, x_N) \\ & \quad = \begin{cases} 1/2^{N-1} & \text{if } \bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N, \\ 0 & \text{else.} \end{cases}\end{aligned} \tag{11}$$

Considering $n = 2$, $f_j = X_j^1 \oplus X_j^2 \oplus \cdots \oplus X_j^{N-1}$, and the direct extension of the protocol described in Fig. 2 for the multipartite case, we see that the communication task is trivialized, implying the maximal violation of the multipartite IC inequality (8).

Additionally, it is important to highlight that the multipartite inequality (8) does not consist of the parallel application of the criterion (1) between each sender with the receiver, an approach followed by the authors of Ref. [21]. Indeed, looking at the simplest tripartite case for $n = 2$, when the receiver Charlie perfectly computes $g_1 = x_1^1 \oplus x_1^2$ and $g_2 = x_2^1 \oplus x_2^2$ all informational terms in the left-hand side of Eq. (1) vanish, showing that the postquantum behavior reached with the protocol of Fig. 2 cannot be detected by this previous approach based on the parallelization of the bipartite IC criterion.

IV. CONCATENATION PROCEDURE

As previously discussed, the first proposal for information causality [9] with the criterion (1) was able to witness the postquantum nature of all nonsignaling correlations beyond Tsirelson's bound. For that, however, it was essential to consider a concatenation procedure involving many copies of the correlation under test. Here we show how such a concatenation can be constructed for the tripartite scenario, also generalizing it to arbitrary multipartite scenarios.

Similarly to the bipartite scenario, the success probability for the protocol in Fig. 2 can be connected to the probability of the resource shared between the parts, more specifically, to the probability $p(a \oplus b \oplus c = xz \oplus yz | x, y, z)$. Clearly, the probabilities of Charlie correctly computing the values of $x_1 \oplus y_1$ and $x_2 \oplus y_2$ are, respectively,

$$\begin{aligned}P_I = {} & \tfrac{1}{4}[p(a \oplus b \oplus c = 0|0,0,0) + p(a \oplus b \oplus c = 0|0,1,0) \\ & + p(a \oplus b \oplus c = 0|1,0,0) + p(a \oplus b \oplus c = 0|1,1,0)],\end{aligned} \tag{12a}$$

$$\begin{aligned}P_{II} = {} & \tfrac{1}{4}[p(a \oplus b \oplus c = 0|0,0,1) + p(a \oplus b \oplus c = 1|0,1,1) \\ & + p(a \oplus b \oplus c = 1|1,0,1) + p(a \oplus b \oplus c = 0|1,1,1)].\end{aligned} \tag{12b}$$

In the particular case where the parties share the correlation described by Eq. (9), we obtain $P_I = P_{II} = 1$.

In Fig. 3, we specify the concatenation procedure for the tripartite communication protocol of Fig. 2. In this case, Alice and Bob initially receive the respective bit-strings $\mathbf{x}$ and $\mathbf{y}$ of length $n = 2^K$ and share $2^K - 1$ identical copies of binary-input and binary-output nonsignaling boxes with Charlie. The success probability that Charlie produces a guess g_j correctly is given by (see Appendix A)

$$p(g_j = x_j \oplus y_j) = \tfrac{1}{2}\left(1 + E_I^{K-r} E_{II}^{r}\right), \tag{13}$$

where r denotes the number of times that Charlie measures $z = 1$ in the K levels of the concatenation code displayed in Fig. 3 and $E_i = 2P_i - 1$ [see Eq. (12)]. By considering this success probability, we show in Appendix C that information causality is always violated when $E_I^2 + E_{II}^2 > 1$. In other words, when combined with a concatenation procedure and multiple copies of the behavior under test, the tripartite information causality inequality (10) leads to a generalization of the bipartite inequality (3), given by

$$E_I^2 + E_{II}^2 \leqslant 1. \tag{14}$$

Similarly to Eq. (10), the multiple copies criterion (14) is maximally violated by the behavior (9) since, for this case, $E_I = E_{II} = 1$. Moreover, for isotropic correlations described by a visibility parameter E and such that $E_I = E_{II} = E$, the tripartite multiple copies inequality is violated when $E > 1/\sqrt{2}$, which is exactly the same bound obtained by the authors of Ref. [9] for the bipartite scenario. However, for the tripartite scenario, the Navascués-Pironio-Acin (NPA) hierarchy [22]

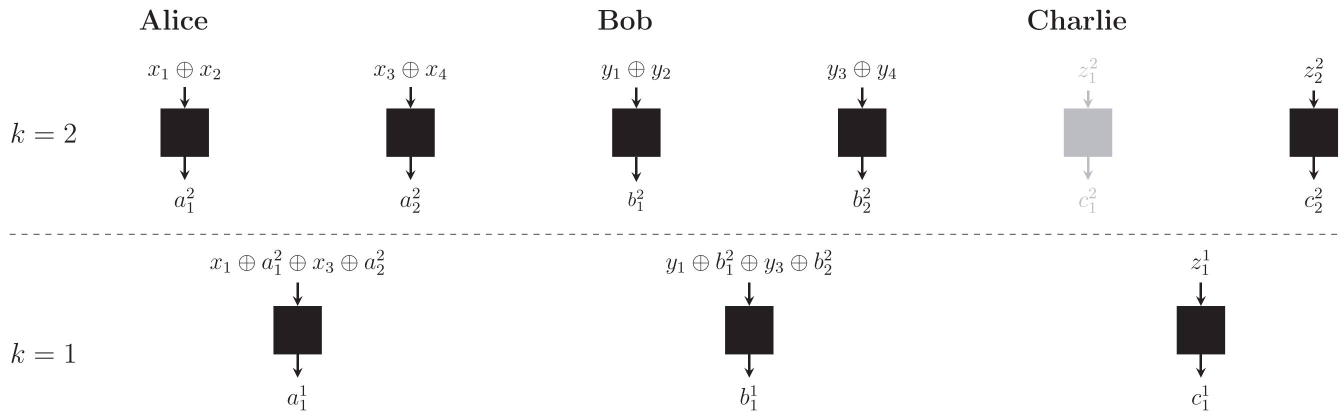

FIG. 3. Concatenation performed by Alice, Bob, and Charlie of the protocol in Fig. 2. Alice and Bob initially receive $n = 2^K$ bits and Charlie receives a K bit string $\{z_1, z_2, \dots, z_K\}$, which indicates which pair $x_j \oplus y_j$ he is interested in, $j = \sum_{l=1}^{K} z_l 2^{l-1}$. Thus, Alice and Bob encode their bits in pairs, just following the protocol in Fig. 2. Now, instead of sending each respective message, they encode pairs of these in other identical NS boxes with the same strategy. So both Alice and Bob perform this procedure until one message remains. Alice and Bob are then allowed to send these one-bit messages to Charlie, who receives the message and to each NS box performs the decoding protocol just as Fig. 2. The idea is that, in a given concatenation level k, Charlie recovers the sum of Alice and Bob messages previously encoded in the current box, which is associated with a subsequently higher level $k+1$ NS box. The picture shows a particular case with $n = 4$, where a_i^k, b_i^k, and c_i^k represent the output of the box i in the level k to Alice, Bob and Charlie, respectively. In the level $k = 1$ Charlie recovers the messages associated with the box $i = 1$ of the level $k = 2$ and is able to recover $x_3 \oplus y_3$ or $x_4 \oplus y_4$ depending on z_2^2.

implies that for any $E \geqslant 1/2$ the corresponding correlation will have a postquantum nature. That is, the tripartite information causality, at least with the specific concatenation considered here, is unable to recover the Tsirelson's bound.

As previously mentioned, the bipartite version of Eq. (14) is equivalent to the quadratic constraint obtained by Uffink [16]. However, for more than two parts, such equivalence no longer holds. For the tripartite scenario, the Uffink inequality reads as

$$(C_{001} + C_{010} + C_{100} - C_{111})^2 + (C_{110} + C_{101} + C_{011} - C_{000})^2 \leqslant 16, \qquad (15)$$

where $C_{xyz} = \sum_{a,b,c} (-1)^{a+b+c} p(a, b, c|x, y, z)$. Indeed, there is no way to alternate between the inequalities (14) and (15) by changing labels. Even more importantly, as we will show in the next section, there are postquantum correlations violating the multiple copies inequality (14) that do not violate the tripartite Uffink inequality (15) (and all the inequalities that are obtained from it by relabeling of parties, measurements, and outcomes).

V. NO CONCATENATION VERSION

The motivation for the generalization of the IC inequality given by Eq. (7) comes from the fact that the upper bound in the information gain of the receiver in Eq. (1) should be understood as the single use of a noiseless classical channel of capacity $|M|$. Interestingly, our results in the multipartite scenario can also take into account this insight. Indeed, for the multipartite scenario described in Sec. III we may consider that each of the $N-1$ senders performs a single use of a classical noisy channel of capacity $C_k = I(M_k : M_k')$, where M_k is the message encoded by each sender k and M_k' is the respective message reaching the decoder after the message passes through the noisy channel. In this case, the criterion defined in Eq. (8) is easily rewritten in terms of the channel capacity by considering data processing inequalities and the fact that M_k completely determines M_k' [that is, M_k' is conditionally independent of any random variable V of the causal structure given M_k, $I(M_k' : V|M_k) = 0$]. As proven in Appendix D, it follows that

$$\sum_{k}^{N-1} \sum_{i}^{n} I\big(X_i^k : X_i^1, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) \leqslant \sum_{k}^{N-1} C_k + \sum_{i}^{n} I\big(X_{i+1}^k, \dots, X_n^k : X_i^k\big). \qquad (16)$$

Particularizing for the tripartite scenario and, for simplicity, assuming completely uncorrelated initial bits, such a result reads as

$$\mathcal{I} \equiv \sum_{i=1}^{n} [I(X_i : Y_i, G_i) + I(Y_i : X_i, G_i)] \leqslant I(M_x : M_x') + I(M_y : M_y'), \qquad (17)$$

where the two senders have initially $\mathbf{x}^1 = (X_1^k, X_2^k, \dots, X_n^k)$ and $\mathbf{x}^2 = (Y_1^k, Y_2^k, \dots, Y_n^k)$, and M_x' and M_y' are the messages reaching the receiver after M_x and M_y pass through the noisy channel, respectively. To illustrate, an application of the noisy tripartite IC inequality will be shown in the next section.

VI. NUMERICAL TESTS

More importantly, to understand the strength of the criteria derived, we considered the following slice of the nonsignaling set

$$p(a, b, c|x, y, z) = \gamma p_{45} + \epsilon p_D + (1 - \gamma - \epsilon) p_W, \qquad (18)$$

where $\gamma, \epsilon \in [0, 1]$, $p_{45}(a, b, c|x, y, z)$ is the distribution we defined in Sec. III in Eq. (9), $p_D(a, b, c|x, y, z) = \delta_{a,0}, \delta_{b,0}\delta_{c,0}$

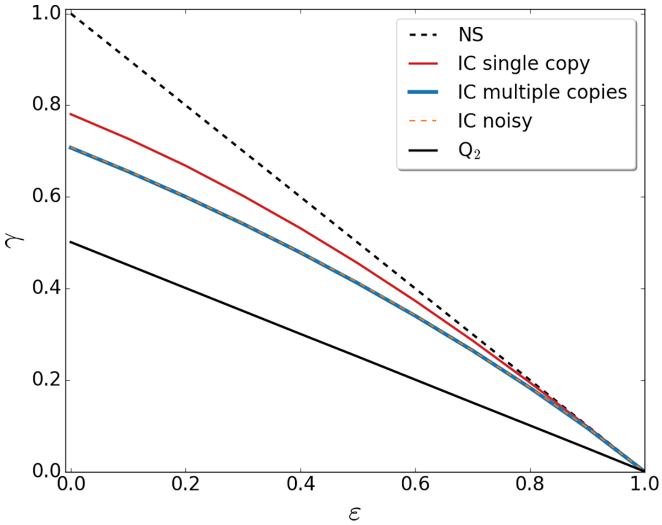

FIG. 4. NS polytope slice given by Eq. (18). Every dot above these curves violates the respective criterion represented. The black dashed and solid lines describe the NS and quantum edges, respectively (the last was computed with the level 2 of NPA hierarchy [22]). The single and multiple copies limits defined by the criteria in Eqs. (10) and (14) are, respectively, depicted with the red and blue solid lines. Finally, the edge defined by the noisy channel criterion (17) is described by the orange dashed line.

and $p_W(a, b, c|x, y, z) = 1/8$. Thus, we obtained Fig. 4, which highlights that Eq. (14) excludes even more supraquantum correlation than Eq. (10). In addition, despite the distance evidenced between the quantum set and IC, we enforced that the bound in Eq. (14) follows from the particular communication protocol depicted in Fig. 3. Therefore, it does not exclude the existence of better protocols, able to single out the quantum set for this slice of the nonsignaling set or rule out postquantum extremal correlations.

In Fig. 4, also we present the edge implied by Eq. (17) for the same slice in Eq. (18), where we consider that all communication is made through a binary symmetric channel that flips the bit with probability ϵ. In this case, to obtain the curve, we followed the results from Ref. [15] and considered $\epsilon \to 1/2$. From this result, it is clear that our stronger criterion in Eq. (14) and this noisy channel approach are in complete agreement, even considering the simplest noisy channel. The codes related to Fig. 4 are available in Ref. [23].

For the tripartite scenario with binary-input and binary-output there exist, 53 856 nonsignaling extremal correlations that are divided into 46 different equivalence classes, among which 45 are supraquantum ones [20]. An important result from Ref. [14] states that class 4 of these could never have its postquantumness detected by principles with a strict bipartite formulation, just as those in Eqs. (1) and (6). Thus, we also checked the ability of Eq. (14) to exclude correlations from class 4, and even more generally, we tested all the 45 supraquantum extremal distributions of the nonsignaling set. In Table I, we highlight those classes for which we could find a violation of Eq. (14). Despite the fact that class 4 does not violate Eq. (14), we stress that this result is limited to a specific protocol performed among the parties. As it always happen with applications of information causality, it is an open problem whether there are other protocols that would ensue that imply a violation of Eq. (10) for class 4 and others of supraquantum correlations. Furthermore, Table I contains the same analysis for the tripartite Uffink inequality (15). From these results, it is clear the no equivalence between the multiple copies IC inequality (14) and the Uffink result (15) since there exist extremal nonsignaling correlations which respect one constraint while violating the other. The codes related to the results from Table I are available in Ref. [24].

TABLE I. Classes of nonsignaling extremal correlations defined in Ref. [20] that violate Eqs. (14) or (15).

Inequality	Extremal boxes
Eq. (14)	35, 37, 38, 40, 41, 42, 43, 44, 45
Eq. (15)	21, 22, 30, 34, 36, 39, 41, 44, 46

VII. CONCLUSION

We proposed a different multipartite communication task in which the previous IC formulation does not detect nonlocal advantage. Thus, by employing the quantum causal structure formalism, we proposed a different criterion to describe IC in such a context and proved its truthfulness for the entire set of quantum correlations, for any number of parts. Furthermore, we proved that our model allows the concatenation approach from Ref. [9], which enabled us to derive even stronger constraints for the multipartite nonsignaling correlations set. In that case, our multipartite inequality proved to be strictly stronger than the multipartite Uffink's inequality from Ref. [16], which contrasts with the previous bipartite result from Ref. [9]. In addition, our findings are in complete agreement with the recent noisy channel approach from Ref. [15], which allows many other analyses for such a multipartite context.

Despite the negative result to single out the class 4 correlations defined in Ref. [20], we emphasize that our results are limited by one specific protocol, which is optimal to Eq. (11), however, it does not ensure that it is optimal for all nonsignaling correlations. Thus, searching for better protocols for different correlations may yield stronger results and it is indeed one of the main interesting further directions. Furthermore, the analysis of nondichotomy scenarios, or even cases where the sender's initial bits are correlated, may also produce interesting results, as previously analyzed in Ref. [25]. Moreover, our findings open a class of nonsequential multipartite RACs, where multiple parts send messages to each other with the task to compute a boolean function of the senders' initial bits. The figure of merit, in this case, is to compute the success probability concerning the receiver to compute correctly such a function. Thus, investigating these new thresholds for such a probability of success may have important implications for quantum information processing.

ACKNOWLEDGMENTS

We thank Marcelo Terra Cunha and Pedro Lauand for fruitful discussions and suggestions. We especially acknowledge Pedro Lauand for providing the extremal nonsignaling behaviors leading to the results in Table I. This work was also supported by the Brazilian National Council for Scientific and Technological Development (CNPq, Grant No. 307295/2020-6), the National Institute for Science and Technology on Quantum Information (INCT-IQ) (Grant No. 465469/2014-0), the Serrapilheira Institute (Grant No. Serra-1708-15763), and the São Paulo Research Foundation FAPESP (Grants No. 2018/07258-7 and No. 2019/00451-9).

APPENDIX A: CONCATENATION IN A MULTIPARTITE COMMUNICATION TASK

Here we extend the tripartite communication task from Sec. III to a general multipartite scenario. Thus, consider N parts, among which $N-1$ are senders that initially have their respective bit-strings $\mathbf{x}^k = (X_1^k, X_2^k, \ldots, X_n^k)$, where $k \in \{1, 2, \ldots, N-1\}$. Each sender encodes a classical message M_k of size $m < n$ to the Nth part, the receiver. This last one needs rightly compute one of n possible initial bits functions $f_j(X_j^1, X_j^2, \ldots, X_j^{N-1})$, by producing the guess G_j, where $j \in \{1, \ldots, n\}$. Just as in the main text, in addition to the classical messages, nonsignaling correlations are allowed among all N parts.

Now consider a little more particular case, where $n = 2$ and $f_j = X_j^1 \oplus X_j^2 \oplus \cdots \oplus X_j^{N-1}$. Just as in the previously described tripartite scenario, we find such a particular multipartite communication task is trivialized by a generalization of the correlation (9) for the $(N, 2, 2)$ Bell scenario, i.e.,

$$p(a_1, a_2, \ldots, a_N | x_1, x_2, \ldots, x_N) = \begin{cases} 1/2^{N-1} & \text{if } \bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N, \\ 0 & \text{else.} \end{cases} \tag{A1}$$

where a_k and x_k, respectively, denote the output and input of the part k. To see this, consider that the N parts perform the strategy depicted in Fig. 2. That is, each sender performs the encoding $x_k = X_1^k \oplus X_2^k$ and $M_k = X_1^k \oplus a_k$ and the receiver computes the guess $G_j = \bigoplus_{k=1}^{N-1} M_k \oplus a_N$. In this case, by considering Eq. (11) we find

$$\begin{aligned} G_j &= \bigoplus_{k=1}^{N-1} (X_1^k \oplus a_k) \oplus a_N \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k\right) \oplus \left(\bigoplus_{k=1}^{N} a_k\right) \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k\right) \oplus \left(\bigoplus_{k=1}^{N-1} x_k x_N\right) \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k\right) \oplus \left(\bigoplus_{k=1}^{N-1} (X_1^k \oplus X_2^k) x_N\right). \end{aligned} \tag{A2}$$

Therefore, if the receiver chooses his measurement as $x_N = j$, when $j = 0$ we have $G_0 = X_1^1 \oplus X_1^2 \oplus \cdots \oplus X_1^{N-1}$, and for $j = 1$ we obtain $G_1 = X_2^1 \oplus X_2^2 \oplus \cdots \oplus X_2^{N-1}$. i.e., the receiver always computes the functions perfectly and trivializes the communication task. It is clear that the task success is related to the probability of the nonsignaling boxes working just as Eq. (11), i.e., $p(a_1 \oplus a_2 \oplus \cdots \oplus a_N = x_1 x_N \oplus x_2 x_N \oplus \cdots \oplus x_{N-1} x_N | x_1, x_2, \cdots, x_N)$. Thus, the probabilities that the receiver computes the function values f_1 and f_2 correctly are, respectively, given by

$$P_I = \frac{1}{2^{N-1}} \left[\sum_{x_1, \ldots, x_{N-1}} p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N | x_1, \ldots, x_{N-1}, x_N = 0 \right) \right], \tag{A3a}$$

$$P_{II} = \frac{1}{2^{N-1}} \left[\sum_{x_1, \ldots, x_{N-1}} p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N | x_1, \ldots, x_{N-1}, x_N = 1 \right) \right]. \tag{A3b}$$

When the parts share Eq. (11), we have $P_I = P_{II} = 1$. However, by introducing a parameter $E \in [0, 1]$, we can investigate other nonsignaling behaviors by means of the following probability of success:

$$p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N \right) = \frac{1}{2}(1 + E). \tag{A4}$$

The perfect correlations of behavior in Eq. (11) are retrieved when $E = 1$ and uniform probabilities are retrieved when $E = 0$.

From this example, one can see that the concatenation approach, depicted in Fig. 3, can also be employed in this multipartite scenario. This is due to the fact that, to complete the task, it is sufficient for the receiver to know only $\bigoplus_{k=1}^{N-1} M_k$, instead of each message M_k. For instance, when $n = 4$, the senders can divide their bits into two pairs and perform the encoding just as in the previous strategy. Now, if instead of sending their respective messages, M_k^0 and M_k^1, the parts encode them in a third NS box (11) by employing Eq. (A2), the receiver is able to recover perfectly one of the functions $\bigoplus_{k=1}^{N-1} M_k^{i=0,1}$. This allows the parts to perform the same decoding one more time, resulting in perfect access by the receiver to one of the functions $f_0 = X_0^1 \oplus X_0^2 \oplus \cdots \oplus X_0^{N-1}$, $f_1 = X_1^1 \oplus X_1^2 \oplus \cdots \oplus X_1^{N-1}$, $f_2 = X_2^1 \oplus X_2^2 \oplus \cdots \oplus X_2^{N-1}$, or $f_3 = X_3^1 \oplus X_3^2 \oplus \cdots \oplus X_3^{N-1}$.

In the most general scenario, the receivers have, initially, $n = 2^K$ bits, share $n-1$ perfect copies of the nonsignaling resource (11), and the senders and the receiver perform the strategy just as depicted in Fig. 3. Here, for each part k, we denote the output and input of the box i of the level l by $a_k^{i,l}$ and $x_k^{i,l}$, respectively. Thus, we may write the guess produced by the receiver as

$$G_j = \left(\bigoplus_{k=0}^{N-1} M_k\right) \oplus \left(\bigoplus_{l=0}^{K-1} a_N^{i_l,l}\right), \tag{A5}$$

where the box i_l is defined in terms of the box measured in the previous level, $i_l = 2i_{l-1} + z_l + 1$, when $l \geqslant 1$. In this case, the receiver performs measurements in K boxes, one in each level, among which $(N-r)$ are to $z_N^{i,l} = 0$ and r to $z_N^{i,l} = 1$, where $r = z_0 + z_1 + \cdots + z_{K-1}$. Just as in the single-copy scenario, the task success is directly related to the probability that the $n-1$ nonsignaling boxes behave as Eq. (11), i.e., Eq. (A4). Thus, when $E < 1$, for each box, there exists a probability that the receiver output $a_{N-1}^{i,l}$ is wrong and the property $\bigoplus_{k=1}^{N-1} a_k^{i,l} = \bigoplus_{k=1}^{N-2} x_k^{i,l} x_N^{i,l}$ does not hold. However, if an even number of mistakes is produced in the outputs of the receiver, then they all cancel each other and the produced guess with Eq. (A5) will be correct. Therefore, the success probability for the multipartite task with concatenation is equal to the probability that the receiver produces an even number of wrong outputs, i.e.,

$$\begin{aligned} p\left(G_j = \bigoplus_{k=0}^{N-2} X_j^k\right) &= Q_{\text{even}}^{(K-r)}(P_I) Q_{\text{even}}^{(r)}(P_{II}) \\ &\quad + Q_{\text{odd}}^{(K-r)}(P_I) Q_{\text{odd}}^{(r)}(P_{II}), \end{aligned} \tag{A6}$$

where P_I and P_{II} are defined in Eq. (A3) and $Q_{\text{even}}^{(s)}(P)$ and $Q_{\text{odd}}^{(s)}(P)$ are given by

$$Q_{\text{even}}^{s}(P) = \sum_{j=0}^{\lfloor \frac{s}{2} \rfloor} \binom{s}{2j} (1-P)^{2j} P^{s-2j} = \frac{1}{2}[1 + (2P-1)^s], \tag{A7a}$$

$$\begin{aligned} Q_{\text{odd}}^{s}(P) &= \sum_{j=0}^{\lfloor \frac{s-1}{2} \rfloor} \binom{s}{2j+1} (1-P)^{2j+1} P^{s-2j-1} \\ &= \frac{1}{2}[1 - (2P-1)^s]. \end{aligned} \tag{A7b}$$

These describe the probabilities of the receiver producing an even and an odd number of mistakes, respectively, after s measurements; P denotes the probability of obtaining the right output in a NS box.

By inserting Eq. (A7) in Eq. (A6) and considering the bias from Eq. (A4) in the probabilities from Eq. (A3), we find the communication task success probability

$$p\left(G_j = \bigoplus_{k=0}^{N-2} X_j^k\right) = \frac{1}{2}\left(1 + E_I^{K-r} E_{II}^{r}\right), \tag{A8}$$

where $E_i = 2P_i - 1$.

APPENDIX B: PROVING NEW IC CRITERIA

In this Appendix, we prove criterion (8), however, for the even more general multipartite scenario described in Appendix A. The strategy will be similar to the one employed in the first bipartite proposal [9], so, for completeness, we start by defining the following mutual information chain rule and data processing inequalities:

$$I(A:B|C) = I(A:B,C) - I(A:C), \tag{B1}$$

$$I(A:B') \leqslant I(A:B), \quad \text{where} \quad B \longrightarrow B'. \tag{B2}$$

Following the description given in Appendix A, first, we consider the following quantity, $I(\mathbf{x}^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c)$, and prove that it is lower bounded by the left-hand side of Eq. (8). By applying the chain rule in Eq. (B1) two times, we obtain

$$\begin{aligned} I(\mathbf{x}^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c) &= I\big(X_1^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big) \\ &\quad + I\big(X_2^k, \dots, X_n^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c | X_1^k\big) \\ &= I\big(X_1^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big) + I\big(X_2^k, \dots, X_n^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \\ &\quad \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c, X_1^k\big) - I\big(X_2^k, \dots, X_n^k : X_1^k\big), \end{aligned} \tag{B3}$$

From data processing (B2) we have

$$I\big(X_2^k, \dots, X_n^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c, X_1^k\big) \geqslant I\big(X_2^k, \dots, X_n^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big). \tag{B4}$$

Furthermore, by applying the chain rule in the first term in the right-hand side of Eq. (B3), and using strong subadditivity $I(A:B|C) \geqslant 0$, we obtain

$$\begin{aligned} I\big(X_1^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big) &= I\big(X_1^k : X_2^1, \dots, X_n^1 | X_1^1, \mathbf{x}^2, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big) \\ &\quad + I\big(X_1^k : X_1^1, \mathbf{x}^2, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big) \\ &\geqslant I\big(X_1^k : X_1^1, \mathbf{x}^2, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k, c\big). \end{aligned} \tag{B5}$$

Therefore, back to Eq. (B3), we write

$$I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\geqslant I\big(X_1^k:X_1^1,\mathbf{x}^2,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c\big)+I\big(X_2^k,\dots,X_n^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c\big)-I\big(X_2^k,\dots,X_n^k:X_1^k\big). \tag{B6}$$

Similarly to Eq. (B5), we can employ the chain rule and strong subadditivity $N-3$ times in the first right-hand side term in Eq. (B6) to highlight only the first bit X_1^k of each bit-string $\mathbf{x}^k$:

$$I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\geqslant I\big(X_1^k:X_1^1,X_1^2,\dots,X_1^{k-1},X_1^{k+1},\dots,X_1^{N-1},M_k,c\big)+I\big(X_2^k,\dots,X_n^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c\big)-I\big(X_2^k,\dots,X_n^k:X_1^k\big). \tag{B7}$$

Notice that the right-side third term in Eq. (B7) is, exactly, $I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)$, but without X_1^k of the bit-string x^k. Therefore, by performing the same steps $n-1$ times, we achieve

$$I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\geqslant\sum_i^n I\big(X_i^k:X_i^1,X_i^2,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},M_k,c\big)-\sum_i^n I\big(X_{i+1}^k,\dots,X_n^k:X_i^k\big). \tag{B8}$$

From the data processing inequality (B2), we write

$$I\big(X_i^k:X_i^1,X_i^2,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},M_k,c\big)\geqslant I\big(X_i^k:X_i^1,X_i^2,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},G_i\big), \tag{B9}$$

and, finally, obtain the lower bound

$$I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\geqslant\sum_i^n I\big(X_i^k:X_i^1,X_i^2,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},G_i\big)-\sum_i^n I\big(X_{i+1}^k,\dots,X_n^k:X_i^k\big). \tag{B10}$$

The next step will be to prove that $I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\leqslant H(M_k)$. So

$$\begin{aligned}I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)&=I(\mathbf{x}^k:M_k|\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},c)\\&\quad+I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},c)\\&=I(M_k:\mathbf{x}^1,\dots,\mathbf{x}^{N-1},c)-I(M_k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},c)\\&\leqslant I(M_k:\mathbf{x}^1,\dots,\mathbf{x}^{N-1},c),\end{aligned} \tag{B11}$$

where here we apply the chain rule two times, considering the nonsignaling between the N parts and the nonnegativity of the mutual information $I(A:B)\geqslant 0$. At this point, just as argued in Ref. [9], from the data processing inequality, we have $I(M_k:\mathbf{x}^1,\dots,\mathbf{x}^{N-1},c)\leqslant I(M_k:M_k)=H(M_k)$, which finally yields

$$I(\mathbf{x}^k:\mathbf{x}^1,\dots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\dots,\mathbf{x}^{N-1},M_k,c)\leqslant H(M_k). \tag{B12}$$

Now, we can put Eqs. (B10) and (B12) together to achieve

$$\sum_i^n I\big(X_i^k:X_i^1,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},G_i\big)\leqslant H(M_k)+\sum_i^n I\big(X_{i+1}^k,\dots,X_n^k:X_i^k\big). \tag{B13}$$

Finally, we recover Eq. (8) by summing inequality (B13) over k and considering nonsignaling between the N parts, i.e., $\sum_k^{N-1}H(M_k)=H(M_1,\dots,M_{N-1})$:

$$\sum_k^{N-1}\sum_i^n I\big(X_i^k:X_i^1,\dots,X_i^{k-1},X_i^{k+1},\dots,X_i^{N-1},G_i\big)\leqslant H(M_1,\dots,M_{N-1})+\sum_k^{N-1}\sum_i^n I\big(X_{i+1}^k,\dots,X_n^k:X_i^k\big). \tag{B14}$$

APPENDIX C: MULTIPLE COPIES INEQUALITY

Here we prove the multipartite generalization of the multiple copies criterion (14), first derived in Ref. [9] for a strict bipartite scenario.

First of all, we need to prove a simplified lower bound for Eq. (8). So, rewriting the left-hand side summation argument in Eq. (8), we have

$$\begin{aligned} I\big(X_i^k : X_i^1, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) &= H\big(X_i^k\big) - H\big(X_i^k|X_i^1, X_i^2 \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) \\ &= 1 - H\big(X_i^k \oplus X_i^1|X_i^1, X_i^2, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) \\ &\geqslant 1 - H\big(X_i^k \oplus X_i^1|X_i^2, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big). \end{aligned} \tag{C1}$$

Here, we particularized to the case where every bit X_i^k is associated with a uniform distribution, $H(X_i^k) = 1$. Further, we considered the fact that $H(A|B, C) = H(A \oplus B|B, C)$ because knowing B results in the same uncertainty about A and $A \oplus B$, and $H(A \oplus B|B, C) \geqslant H(A \oplus B|C)$, i.e., to remove the conditioning in B does not increase the uncertainty of $A \oplus B$. This same argument can be applied $N - 2$ times to move every conditioned random variable in the right-hand side of Eq. (C1):

$$I\big(X_i^k : X_i^1, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) \geqslant 1 - H\big(X_i^1 \oplus X_i^2 \oplus \dots \oplus X_i^{N-1} \oplus G_i\big). \tag{C2}$$

However, from the communication task, when $X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 0$, we necessarily have $G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}$. Thus, the probability $p(X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 0)$ is exactly the success probability of the receiver, $p(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1})$, while $p(X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 1)$ is the complementary part. Therefore, the right-hand side term from Eq. (C2) can be written in terms of the binary entropy, which in Eq. (B14) finally yields

$$(N-1)\sum_i^n \big\{1 - h\big[p\big(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}\big)\big]\big\} \leqslant \mathcal{I} \leqslant H(M_1, \dots, M_{N-1}). \tag{C3}$$

Notice that we considered the fact that the left-hand side has no dependence on the index k. Furthermore, the rightmost term in Eq. (B14) does not appear in Eq. (C3) because we are assuming a uniform distribution for every initial bit X_i^k.

At this point, we particularize our description to the concatenation strategy earlier described in Appendix A. Here we rewrite the left-hand side summation in Eq. (C3) in terms of the number of instances r where the receiver performed measurement $x_{n_j^k} = 1$ and substitute the concatenation success probability (A8)

$$\begin{aligned} (N-1)\sum_i^n \big\{1 - h\big[p\big(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}\big)\big]\big\} &= (N-1)\sum_r^K \binom{K}{r}\left[1 - h\left(\frac{1 + E_I^{K-r}E_{II}^r}{2}\right)\right] \\ &\geqslant \frac{(N-1)}{2\ln 2}\sum_r^K \binom{K}{r}\big(E_I^2\big)^{N-r}\big(E_{II}^2\big)^r \\ &= \frac{(N-1)}{2\ln 2}\big(E_I^2 + E_{II}^2\big)^K, \end{aligned} \tag{C4}$$

where we consider $1 - h(\frac{1+y}{2}) \geqslant \frac{y^2}{2\ln 2}$ and $E_i = 2P_i - 1$, from Eq. (A3). After performing such encoding, each sender sends only a single bit message. Thus, $H(M_1, \dots, M_{N-1})$ in Eq. (C3) is always fixed in $N - 1$, necessarily. Therefore, with Eqs. (C3) and (C4), we find that, when $E_I^2 + E_{II}^2 > 1$, the proposed criterion (8) can always be violated by some concatenation protocol with K levels. Thus, we finally conclude the proof for the previously mentioned criterion in Eq. (14)

$$E_I^2 + E_{II}^2 \leqslant 1. \tag{C5}$$

APPENDIX D: NEW INEQUALITY IN TERMS OF NOISY CHANNEL CAPACITY

Here we prove the inequality (16), where the senders communicate their messages M_k through a single use of a noisy channel to the receiver. The proof for the noiseless version (8) is essentially valid in this context, but it is necessary to introduce a new variable M_k', representing the message after the action of the channel, on step Eq. (B11) to obtain the upper bound in terms of the channel capacity $C_k = I(M_k : M_k')$. So, we have

$$\begin{aligned} &I(\mathbf{x}^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k', c) \\ &\quad\leqslant I(M_k' : \mathbf{x}^1, \dots, \mathbf{x}^{N-1}, c) \\ &\quad\leqslant I(M_k' : \mathbf{x}^1, \dots, \mathbf{x}^{N-1}, c, M_k), \end{aligned} \tag{D1}$$

where we considered the data processing inequality in the second step. As mentioned in the main text, M_k completely determines M_k', so M_k' is conditionally independent of any random variable from the causal structure, i.e., $I(M_k' : V|M_k) = 0$. Thus, we may write $I(M_k' : \mathbf{x}^1, \dots, \mathbf{x}^{N-1}, c, M_k) - I(M_k' : M_k) = 0$ and obtain in Eq. (D1)

$$\begin{aligned} &I(\mathbf{x}^k : \mathbf{x}^1, \dots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \dots, \mathbf{x}^{N-1}, M_k', c) \\ &\quad\leqslant I(M_k' : M_k) = C_k. \end{aligned} \tag{D2}$$

The next steps are quite similar as Appendix B, therefore, we may write

$$\begin{aligned} &\sum_k^{N-1}\sum_i^n I\big(X_i^k : X_i^1, \dots, X_i^{k-1}, X_i^{k+1}, \dots, X_i^{N-1}, G_i\big) \\ &\quad\leqslant \sum_k^{N-1} C_k + \sum_i^n I\big(X_{i+1}^k, \dots, X_n^k : X_i^k\big). \end{aligned} \tag{D3}$$

[1] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] B. S. Cirel'son, Quantum generalizations of bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[4] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom., Found. Phys. **24**, 379 (1994).

[5] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial, Phys. Rev. Lett. **96**, 250401 (2006).

[6] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Proc. R. Soc. London, Ser. A **466**, 881 (2009).

[7] T. Fritz, A. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, Nat. Commun. **4**, 2263 (2013).

[8] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, Almost quantum correlations, Nat. Commun. **6**, 6288 (2015).

[9] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature (London) **461**, 1101 (2009).

[10] J. Allcock, N. Brunner, M. Pawlowski, and V. Scarani, Recovering part of the boundary between quantum and nonquantum correlations from information causality, Phys. Rev. A **80**, 040103(R) (2009).

[11] R. Chaves, C. Majenz, and D. Gross, Information–theoretic implications of quantum causal structures, Nat. Commun. **6**, 5766 (2015).

[12] B. Yu and V. Scarani, Information causality beyond the random access code model, arXiv:2201.08986.

[13] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Quantum Correlations Require Multipartite Information Principles, Phys. Rev. Lett. **107**, 210403 (2011).

[14] T. H. Yang, D. Cavalcanti, M. L. Almeida, C. Teo, and V. Scarani, Information-causality and extremal tripartite correlations, New J. Phys. **14**, 013061 (2012).

[15] N. Miklin and M. Pawłowski, Information Causality without Concatenation, Phys. Rev. Lett. **126**, 220403 (2021).

[16] J. Uffink, Quadratic Bell Inequalities as Tests for Multipartite Entanglement, Phys. Rev. Lett. **88**, 230406 (2002).

[17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[18] S. G. A. Brito, M. G. M. Moreno, A. Rai, and R. Chaves, Nonlocality distillation and quantum voids, Phys. Rev. A **100**, 012102 (2019).

[19] A. Rai, C. Duarte, S. Brito, and R. Chaves, Geometry of the quantum set on no-signaling faces, Phys. Rev. A **99**, 032106 (2019).

[20] S. Pironio, J.-D. Bancal, and V. Scarani, Extremal correlations of the tripartite no-signaling polytope, J. Phys. A: Math. Theor. **44**, 065303 (2011).

[21] L.-Y. Hsu, Multipartite information causality, Phys. Rev. A **85**, 032115 (2012).

[22] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).

[23] L. Pollyceno, Code concerning Fig. 4 (2022), https://github.com/Pollyceno/TripartiteSlices.

[24] L. Pollyceno, Code concerning the results from Table I (2022), https://github.com/Pollyceno/TripartiteViolation.

[25] D. Cavalcanti, A. Salles, and V. Scarani, Macroscopically local correlations can violate information causality, Nat. Commun. **1**, 136 (2010).